

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 25, 2003, 22:29:57 ; Search time 2180 Seconds
(without alignments)
7013.092 Million cell updates/sec

Title: US-09-701-023-1_COPY_81_1024
Perfect score: 944
Sequence: 1 tgcgcactctctctcttc.....atttgtcttgaagctttaa 944

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estov:*
5: em_estpl:*
6: em_estro:*
7: em_hic:*
8: gb_est1:*
9: gb_est2:*
10: gb_est3:*
11: gb_est4:*
12: gb_est5:*
13: gb_est6:*
14: gb_est7:*
15: gb_est8:*
16: gb_est9:*
17: gb_est10:*
18: gb_est11:*
19: gb_est12:*
20: gb_est13:*
21: gb_est14:*
22: gb_est15:*
23: gb_est16:*
24: gb_est17:*
25: gb_est18:*
26: gb_est19:*
27: gb_est20:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	312.4	33.1	592	17	B98482 T25H20TRB T
C 2	296.8	31.4	508	17	B67977 T25H20TR TA
C 3	182.2	19.3	641	17	BH472044 BOHUA43TR
C 4	59.4	6.3	775	17	BH472100 BOHUR55TR
C 5	45	4.8	645	10	BE188445
C 6	41	4.3	1672	10	BE622474 601441274

C 7	40.8	4.3	937	17	CNS006XP
C 8	40.6	4.3	863	17	BH154230
C 9	40.4	4.3	552	17	AQ338133
C 10	40.4	4.3	859	17	CNS06CGK
C 11	39	4.1	519	14	BQ515341
C 12	39	4.1	522	13	B1432782
C 13	39	4.1	525	14	BQ515340
C 14	39	4.1	1105	17	CNS07ARO
C 15	39	4.1	1167	12	BG856705
C 16	38.8	4.1	673	14	BQ390257
C 17	38.8	4.1	638	9	AL647711
C 18	38.6	4.1	937	17	CNS006ST
C 19	38.4	4.1	637	17	AZ097834
C 20	38.4	4.1	653	13	B1927000
C 21	38.2	4.0	477	13	B1465229
C 22	38	4.0	487	12	BF097623
C 23	37.8	4.0	565	17	AQ490219
C 24	37.4	4.0	331	12	BF587816
C 25	37.4	4.0	921	17	CNS04NCO
C 26	37.2	3.9	544	9	AL630678
C 27	37.2	3.9	551	14	BQ418553
C 28	37.2	3.9	889	17	B21212
C 29	37.2	3.9	997	17	CNS005TE
C 30	37	3.9	660	17	AG142495
C 31	37	3.9	695	17	AZ938420
C 32	36.8	3.9	275	17	AZ768950
C 33	36.8	3.9	609	13	B1514873
C 34	36.8	3.9	636	10	BB635500
C 35	36.8	3.9	902	13	B1856118
C 36	36.6	3.9	443	17	AQ850356
C 37	36.6	3.9	550	17	AQ847422
C 38	36.6	3.9	609	17	AQ847422
C 39	36.6	3.9	688	17	AZ775117
C 40	36.6	3.9	1578	12	BF575807
C 41	36.4	3.9	238	17	AZ263790
C 42	36.4	3.9	374	12	BE762203
C 43	36.4	3.9	503	17	BH077388
C 44	36.4	3.9	665	17	AQ630761
C 45	36.2	3.8	684	17	AG040879

ALIGNMENTS

RESULT 1
B98482/c

LOCUS
DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

B98482 T25H20TRB TAMU Arabidopsis thaliana genomic clone T25H20, DNA sequence.
B98482 B98482.1 GI:3000561
GSS.
Arabidopsis thaliana
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 592)
Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wible, C., Adams, M.D. and Venter,
J.C.
A BAC End Sequence Database for Identifying Minimal Overlaps in
Arabidopsis Genomic Sequencing. Update 3
Unpublished (1997)
Other_GSSs: T25H20TF T25H20TFB T25H20TR
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13 Reverse

Class: BAC ends
High quality sequence stop: 592.

FEATURES

source
1..592
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="T25H20"
/clone.lib="TAMU"
/sex="hermaphrodite"
/note="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII
; Produced by Rod Wing"
BASE COUNT 195 a 127 c 91 g 176 t 3 others
ORIGIN

Query Match 33.1%; Score 312.4; DB 17; Length 592;
Best Local Similarity 82.1%; Pred. No. 9.9e-82;
Matches 412; Conservative 0; Mismatches 1; Indels 89; Gaps 1;
529 AGAAGAACGTTGGATGGTGATCAGATAATAGTTCGATCCCAACGGTGGTGATTTT 588
|||||
503 AGAAGAACGTTGGATGGTGATCAGATAATAGTTCGATCCCAACGGTGGTGATTTT 444
|||||
509 CGAATACACAAATGATTCCTCTCCGATGACGGCTAGCATGATCTTCTTCAATCAG 648
|||||
443 CGAATACACAAATGATTCCTCTCCGATGACGGCTAGCATGATCTTCTTCAATCAG 384
|||||
649 ATCATCATCAGAGGACCAAGGTTTCCTTTATGATCATAGATCGCTAGACAGCTTCAG 708
|||||
383 ATCATCATCAGAGGACCAAGGTTTCCTTTATGATCATAGATCGCTAGACAGCTTCAG 324
|||||
709 TTCTGCTTCTAGTACTACTATTAATCTCTTATTCACGAGGCAACAATCATATC 763
|||||
323 TTCTGCTTCTAGTACTACTATTAATCTCTTATTCACGAGGCAACAATCATATC 264
|||||
764 ----- 763
263 TAAGTATAGTCCATTTATTAATCTCATATATAGTATATATATATATATATATAT 204
|||||
764 -----GGACCAATGGAGGAATTTGGAGCTACATGGAAGG 799
|||||
203 TATTTGATTTAACTGCTGGGTTTAGGGACCAATGGAGGAATTTGGAGCTACATGGAAGG 144
|||||
800 AAACCCTAGAATGGATCAGAGGTGTGAAGGATGACAGTTTTTCGGGGAATATGG 859
|||||
143 AAACCCTAGAATGGATCAGAGGTGTGAAGGATGACAGTTTTTCGGGGAATATGG 84
|||||
860 TGAAGAGTTTCAGTGGTGCTACAACTGCTCACTCTGATGATGATGATGATGATGATGAT 919
|||||
83 TGAAGAGTTTCAGTGGTGCTAAACGCTGCTCACTCTGATGATGATGATGATGATGATGAT 24
|||||
920 CATTGATTTGCTTGAAGCTT 941
|||||
23 CATTGATTTGCTTGAAGCTT 2

RESULT 2
LOCUS B67977 T25H20TR TAMU Arabidopsis thaliana genomic clone T25H20, DNA
DEFINITION sequence.
ACCESSION B67977
VERSION B67977.1 GI:2666731
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 508)
Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and Venter
,J.C.

TITLE

Arabidopsis Genomic Sequencing. Update 3
Unpublished (1997)
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 508.

FEATURES

source
1..508
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="T25H20"
/clone.lib="TAMU"
/sex="hermaphrodite"
/note="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII
; Produced by Rod Wing"
BASE COUNT 158 a 117 c 82 g 151 t
ORIGIN

Query Match 31.4%; Score 296.8; DB 17; Length 508;
Best Local Similarity 81.6%; Pred. No. 4.1e-77;
Matches 408; Conservative 0; Mismatches 2; Indels 90; Gaps 2;

531 AAGAAACGTTGGATGGTGATCAGATAATAGTTCGATCCCAACGGTGGTGATTTTCG 590
|||||
508 AAGAAACGTTGGATGGTGATCAGATAATAGTTCGATCCCAACGGTGGTGATTTTCG 449
|||||
591 AAATACAAATGATTCCTCTCCGATGAACGGCTACGATCATCTTCTTCAATCAGAT 650
|||||
448 AAATACAAATGATTCCTCTCCGATGAACGGCTACGATCATCTTCTTCAATCAGAT 389
|||||
651 CATCATCAGAGGACCAAGGTTTCCTTTATGATCATAGATCGCTAGACAGCTTCAGTT 710
|||||
388 CATCATCAGAGGACCAAGGTTTCCTTTATGATCATAGATCGCTAGACAGCTTCAGTT 329
|||||
711 TCTGCTTCTAGTACTACTATTAATCTCTTATTCACGAGGCAACAATCATATC 763
|||||
328 TCTGCTTCTAGTACTACTATTAATCATATTTCAACGAGGCAACAATCATATC 269
|||||
764 ----- 763
268 AGTATAGTCCATTTATTAATCTCATATATAGTATATATATATATATATATATATAT 209
|||||
764 -----GGGACCAATGGAGGAATTTGGAGCTACATGGAAGAA 801
|||||
208 TTGATTTAACTGGTGGGTTTAGGGACCAATGGAGGAATTTGGAGCTAGATGGAAGAA 149
|||||
802 ACCCTAGAAATGGATCAGAGGTGTGAAGGATGACGAGTTTTTTT-CGGGGGAATATGGT 860
|||||
148 ACCCTAGAAATGGATCAGAGGTGTGAAGGATGACGAGTTTTTTTACC GGGAATATGGT 89
|||||
861 GAAAGAGTTTCAGTGGTGCTACAACTGCTCACTCTGATGATGATGATGATGATGATGAT 920
|||||
88 GAAAGAGTTTCAGTGGTGCTACAACTGCTCACTCTGATGATGATGATGATGATGATGAT 29
|||||
921 ATTGATTTGCTTGAAGCT 940
|||||
28 ATTGATTTGCTTGAAGCT 9

RESULT 3

BH472044
LOCUS BOHUA43TR BOHU Brassica oleracea genomic clone BOHUA43, DNA
DEFINITION sequence.
ACCESSION BH472044

Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSLFMTSDQNSVGNPNLLNRLVNVSSGEIRTELKSRKPGSKTGQOKKPT 60
 DB 1 MATSLFMTSDQNSVGNPNLLNRLVNVSSGEIRTELKSRKPGSKTGQOKKPT 60

QY 61 LRGMGVAKLERQRIEERKQKLAATVGDTSVASISNNATRLPVPDGVVLOGFPSSLG 120
 DB 61 LRGMGVAKLERQRIEERKQKLAATVGDTSVASISNNATRLPVPDGVVLOGFPSSLG 120

QY 121 SNRIYCGVGSGQVMIDPVIWPGFVETSTTHELSSISNPQMFNASSNNRCDTCFKKR 180
 DB 121 SNRIYCGVGSGQVMIDPVIWPGFVETSTTHELSSISNPQMFNASSNNRCDTCFKKR 180

QY 181 LDGQNNVVRSGGFGSKYTMIPPMNGYDQYLLQSDHHQSGFLYDHRARAASVSAS 240
 DB 181 LDGQNNVVRSGGFGSKYTMIPPMNGYDQYLLQSDHHQSGFLYDHRARAASVSAS 240

QY 241 STTINPYFNATNHTGPMEEFGSYMEGNPRNGSGGVKEYEFFFGKYGERVSVVATTSSLV 300
 DB 241 STTINPYFNATNHTGPMEEFGSYMEGNPRNGSGGVKEYEFFFGKYGERVSVVATTSSLV 300

QY 301 GDCSPNTIDLKSL 314
 DB 301 GDCSPNTIDLKSL 314

RESULT 2
 ID 081836 PRELIMINARY; PRT; 314 AA.
 AC 081836;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE Hypothetical 34.1 kDa protein (NOZLE/SPOROCYTELESS).
 GN M412.140 OR AT4G27330.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Reichert B.J., Barel E., Hoheisel J., Mewes H.W., Mayer K.,
 RA Schueller C., Bevan M.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 [2]
 SEQUENCE FROM N.A.
 Reichert B.J., Barel E., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL030978; CAA19727.1;
 DR EMBL; AL161566; CAB79588.1;
 KW Hypothetical protein.
 SQ SEQUENCE 314 AA; 34078 MW; B0236D910544CEC3 CRC64;

Query Match 99.6%; Score 1628; DB 10; Length 314;
 Best Local Similarity 99.7%; Pred. No. 3.5e-130;
 Matches 313; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MATSLFMTSDQNSVGNPNLLNRLVNVSSGEIRTELKSRKPGSKTGQOKKPT 60
 DB 1 MATSLFMTSDQNSVGNPNLLNRLVNVSSGEIRTELKSRKPGSKTGQOKKPT 60

QY 61 LRGMGVAKLERQRIEERKQKLAATVGDTSVASISNNATRLPVPDGVVLOGFPSSLG 120
 DB 61 LRGMGVAKLERQRIEERKQKLAATVGDTSVASISNNATRLPVPDGVVLOGFPSSLG 120

QY 121 SNRIYCGVGSGQVMIDPVIWPGFVETSTTHELSSISNPQMFNASSNNRCDTCFKKR 180
 DB 121 SNRIYCGVGSGQVMIDPVIWPGFVETSTTHELSSISNPQMFNASSNNRCDTCFKKR 180

Db 121 SNRIYCGVGSGQVMIDPVIWPGFVETSTTHELSSISNPQMFNASSNNRCDTCFKKR 180
 QY 181 LDGQNNVVRSGGFGSKYTMIPPMNGYDQYLLQSDHHQSGFLYDHRARAASVSAS 240
 DB 181 LDGQNNVVRSGGFGSKYTMIPPMNGYDQYLLQSDHHQSGFLYDHRARAASVSAS 240

QY 241 STTINPYFNATNHTGPMEEFGSYMEGNPRNGSGGVKEYEFFFGKYGERVSVVATTSSLV 300
 DB 241 STTINPYFNATNHTGPMEEFGSYMEGNPRNGSGGVKEYEFFFGKYGERVSVVATTSSLV 300

QY 301 GDCSPNTIDLKSL 314
 DB 301 GDCSPNTIDLKSL 314

RESULT 3
 ID 09LWFL PRELIMINARY; PRT; 320 AA.
 AC 09LWFL;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical protein (OSUNBa0038J17.1 protein).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 clone: P0031E09.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
 clone: OSUNBa0038J17.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP002092; BAA96171.1;
 DR EMBL; AP003104; BAB55710.1;
 KW Hypothetical protein.
 SQ SEQUENCE 320 AA; 34068 MW; AC256F3C99245D2B CRC64;

Query Match 8.0%; Score 131; DB 10; Length 320;
 Best Local Similarity 21.1%; Pred. No. 0.0049;
 Matches 56; Conservative 33; Mismatches 78; Indels 98; Gaps 9;

QY 53 OQKQKPTLRGMGVAKLERQRIEERKQKLAATVGDTSVASISNNATRLPVPDGVV 112
 DB 37 KNRKPKVPQRLGVAQLEKLRTEEQKMEGAALSHGALGGDGVGLHLPPLPPPSLAL 96

QY 113 Q-----GFPSSLSGNRIYCGVGSGQVMIDPVIWPGFVETSTTHELSSISNPQMFNASSNNRCDTCFKKR 143
 DB 97 SALQSRSAADGGFPAALWS-----PVDPIKHPYKRSLCPQPPSPMVSTGLSL 143

QY 148 TSTTHELSSISNPQMFNASSNNRCDTCFKKRLDGQNNVVRSGGFGSKYTMIPPMN 207
 DB 144 TASARHPTPEPSN-QMYSNGSGSSSSAAAEER-----GMAAFRTTTSRAPPL- 190

QY 208 GYDQYLLQSDHHQSGFLYDHRARAASVSASTT-----INP-----Y 247
 DB 191 -----VAPFSARTTTTTELGLAGLADVGPDLRYEFRATNY 224

QY 248 FNEATNHTGPMEEFGSYMEGNPRNG 272
 DB 225 FSANANYSDWT SDF-AHCKSKENG 248

RESULT 4

use for a sequence from N. submitted 6/98



sporogenesis and encodes a novel nuclear protein
 Genes Dev. 13 (16), 2108-2117 (1999)
 99396716
 JOURNAL MEDLINE
 10465788
 PUBLISHED
 2 (bases 1 to 1302)
 REFERENCE
 Yang, W.-C., Ye, D., Xu, J. and Sundaresan, V.
 Direct Submission
 TITLE
 Submitted (14-JUN-1999) Plant Molecular Genetics, Institute of
 Molecular Agrobology, 1 Research Link, Singapore 117604, Singapore
 Location/Qualifiers
 1. 1302
 /organism="Arabidopsis thaliana"
 /cultivar="Landsberg erecta"
 /db_xref="taxon:3702"
 /chromosome="4"
 /map="Pgl1-mil23"
 /note="nucleus"
 1. 1302
 /gene="SPL"
 80. 1024
 /gene="SPL"
 /note="MADS-box related protein"
 /codon_start=1
 /product="sporocytelless"
 /protein_id="AA045344.1"
 /db_xref="GI:5566240"
 /translation="MATSLFFMTDQNSVGNPNDLLRNLVNSGEIRTELKSRG
 RKFGSTGQKQKRLRGMVAKLQRLEKEKKOLAAATVGDTSVASISNNATRL
 PVPDVGVLQGFPSLSGSKRIYCGVSGOVMDIPVSPGVETSTTHELSSIN
 POMFNASNNRCDTCFKKRLDGDONNVRSNGGFSKYIWIIPPMNGYDQYLLQSDH
 HORSQFLYDHRHARAASVSASTIINPYNEATNHTGPMEEFGSYMEGNPRNGSGV
 KEYEFPFGKIGERVSVVATSSVLGDCSPNTIDLSLK"

BASE COUNT
 ORIGIN

402 a 273 c 275 g 352 t
 Query Match 100.0%; Score 944; DB 8; Length 1302;
 Best Local Similarity 100.0%; Pred. No. 2.8e-260;
 Matches 944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGGCGACTCTCTCTTCTTCAATGCAACAGATCAAAATCCGTCGGAAACCCAAACGATC 60
 DB 81 TGGCGACTCTCTCTTCTTCAATGCAACAGATCAAAATCCGTCGGAAACCCAAACGATC 140
 QY 61 TTCTGAGAAACACCCGCTTCTCTCAATAGCTCCGGGAGATCCGGACAGACACTGA 120
 DB 141 TTCTGAGAAACACCCGCTTCTCTCAATAGCTCCGGGAGATCCGGACAGACACTGA 200
 QY 121 AGATCTGCTGTCGGAAACCCAGGATCGAAGACAGCTCAGCAAAACAGAGAAACCAACGT 180
 DB 201 AGATCTGCTGTCGGAAACCCAGGATCGAAGACAGCTCAGCAAAACAGAGAAACCAACGT 260
 QY 181 TGAGAGAAATGGGTGTAGCAAGCTCGAGCTCAGAGATCGAAGAAAGAAAGCAAC 240
 DB 261 TGAGAGAAATGGGTGTAGCAAGCTCGAGCTCAGAGATCGAAGAAAGAAAGCAAC 320
 QY 241 TCAGCGCCGCCACAGTCGGAGACAGCTCATAGTAGCATGCTCTTAACAACGCTACCC 300
 DB 321 TCAGCGCCGCCACAGTCGGAGACAGCTCATAGTAGCATGCTCTTAACAACGCTACCC 380
 QY 301 GTTTACCGCTACCGGTAGACCCGGTCTTGTGCTACAGGCTTCCCAAGCTCACTCGGA 360
 DB 381 GTTTACCGCTACCGGTAGACCCGGTCTTGTGCTACAGGCTTCCCAAGCTCACTCGGA 440
 QY 361 GCAACAGAGTCTATTGTGGTGGAGTCGGGTCCAGGTATGATCGACCCGGTTATTT 420
 DB 441 GCAACAGAGTCTATTGTGGTGGAGTCGGGTCCAGGTATGATCGACCCGGTTATTT 500
 QY 421 CTCATGGGGTTTGTGAGACCTCTCCACTACTCATGAGCTCTCTTCAATCTCAATC 480
 DB 501 CTCATGGGGTTTGTGAGACCTCTCCACTACTCATGAGCTCTCTTCAATCTCAATC 560
 QY 481 CTCAAAATGTTAAAGCTTCTTCCAAATAATCGCTGTCACACTTCTTCAAGAAAGACGTT 540

Db 561 CTCAAAATGTTAAAGCTTCTTCCAAATAATCGCTGTCACACTTGCCTCAGAAGAAGACGTT 620
 QY 541 TGGATGGTGATCAGAAATATAGTAGTTCGATCAACCGTGGTGGATTTTCGAAATACACAA 600
 Db 621 TGGATGGTGATCAGAAATATAGTAGTTCGATCAACCGTGGTGGATTTTCGAAATACACAA 680
 QY 601 TGATTCCTCTCCGATGAACGGCTAGCATGATCTTCTTCAATCAGATCATCATCAGA 660
 Db 681 TGATTCCTCTCCGATGAACGGCTAGCATGATCTTCTTCAATCAGATCATCATCAGA 740
 QY 661 GGAGCCAAAGGTTCTTATGATCATAGAAATCGCTAGAGCAGCTTCAGTTTCTGCTTCTA 720
 Db 741 GGAGCCAAAGGTTCTTATGATCATAGAAATCGCTAGAGCAGCTTCAGTTTCTGCTTCTA 800
 QY 721 GTACTACTATTAATCCTTATTTCAACGAGGACAAATATACGAGGACCAATGGAGGAT 780
 Db 801 GTACTACTATTAATCCTTATTTCAACGAGGACAAATATACGAGGACCAATGGAGGAT 860
 QY 781 TTGGAGCTACATGAAGGAAACCTAGAAATGGATGAGGAGTGTGAAGGATGACGAGT 840
 Db 861 TTGGAGCTACATGAAGGAAACCTAGAAATGGATGAGGAGTGTGAAGGATGACGAGT 920
 QY 841 TTTTCCGGGAAATATGTTGAAAGAGTTTCAGTGGTGGCTACAAACGTCGTCACCTCGTAG 900
 Db 921 TTTTCCGGGAAATATGTTGAAAGAGTTTCAGTGGTGGCTACAAACGTCGTCACCTCGTAG 980
 QY 901 GTGATTGACGCTTAATACCATGATTTGCTTCTGCTGAAAGCTTTAA 944
 Db 981 GTGATTGACGCTTAATACCATGATTTGCTTCTGCTGAAAGCTTTAA 1024
 RESULT 2
 AF146794
 LOCUS
 DEFINITION Arabidopsis thaliana NOZZLE (NZZ) gene, complete cds.
 ACCESSION AF146794
 VERSION AF146794.1
 KEYWORDS
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 1133)
 AUTHORS Schiefthaler, U., Balasubramanian, S., Sieber, P., Chevallier, D.,
 Wisman, E. and Schneitz, K.
 TITLE Molecular analysis of NOZZLE, a gene involved in pattern formation
 and early sporogenesis during sex organ development in Arabidopsis
 thaliana
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (20), 11664-11669 (1999)
 MEDLINE 99432290
 PUBLISHED 10500234
 REFERENCE 2 (bases 1 to 1133)
 AUTHORS Schiefthaler, U., Balasubramanian, S., Chevallier, D., Sieber, P. and
 Schneitz, K.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Plant Biology, University of
 Zurich, Zollikonerstr. 107, Zurich, ZH 8008, Switzerland
 FEATURES
 source
 1. 1133
 /organism="Arabidopsis thaliana"
 /cultivar="Landsberg erecta"
 /db_xref="taxon:3702"
 /map="between Pgl1 and PRHA"
 /chromosome="4"
 <1..>1133
 /gene="NZZ"
 join(<1..531,631..864,954..>1133)
 /gene="NZZ"
 /product="NOZZLE"
 join(1..531,631..864,954..1133)
 /gene="NZZ"
 /note="involved in ovule and anther development and early

